

**AMENDMENTS TO THE SPECIFICATION**

Please amend the subtitle beginning on page 1, line 5 as follows:

~~Technical~~ Background of the Invention

Field of the Invention

Please amend the subtitle beginning on page 1, line 11 as follows:

~~Background Art~~ Description of the Related Art

Please amend the paragraph beginning on page 2, line 21 with the following amended paragraph:

However, since the conventional technology needs ~~high~~ large enough memory and a high-performance CPU, it suffers from the problem of being impossible for a terminal such as a mobile terminal, which has only limited memory and a CPU with a limited processing performance, to realize the technology.

Please amend the subtitle beginning on page 3, line 20 as follows:

~~Disclosure~~ Summary of Invention

Please amend the paragraph beginning on page 7, line 14 with the following amended paragraph:

Fig.1 is a block diagram showing an embodiment of an image data distribution system according to the present ~~invention~~ invention;

~~Fig.2 is an illustration~~ Figs. 2A and 2B are illustrations showing layouts of a plurality of cameras for preparing multiple-view image ~~data~~ data;

~~Fig.3 is an illustration~~ Figs. 3A and 3B are illustrations showing the left eye viewpoint L and the right eye viewpoint R for which image data is generated by ~~interpolation~~ interpolation;

~~Fig.4 is an illustration~~Figs. 4A and 4B are illustrations showing areas to be cut out from the generated image data, depending on the resolution of a display ~~unit-unit;~~

Fig.5 is an illustrative view showing a mobile terminal to be a ~~client-client;~~

~~Fig.6 is an illustration~~Figs. 6A, 6B, and 6C are illustrations showing examples of states of stored data of multiple-view image ~~data-data;~~

~~Fig.7 is an illustration~~Figs. 7A, 7B, and 7C are illustrations showing examples of states of generated image ~~data-data;~~

~~Fig.8 is a diagram~~Figs. 8A and 8B are diagrams showing storage and extraction of multiple-view image ~~data-data;~~

Fig.9 is a flowchart showing the sequence of processing on a ~~server-server;~~

Fig.10 is a flowchart showing the sequence of processing on a ~~client-client;~~

Fig.11 is an illustration showing an example of coding multiple viewpoint moving picture data based on ~~MPEG-4-MPEG-4;~~

Fig.12 is a diagram showing multiple viewpoint image data added with management ~~information-information;~~

~~Fig.13 is a chart~~Figs. 13A and 13B are charts showing one example of management ~~information-information;~~

Fig.14 shows an expected connection state between a server and clients in the present ~~embodiment-embodiment;~~

Fig.15 is a flowchart showing details of the processing of an image ~~generator-generator; and~~

Fig.16 is a diagram showing a conventional example.

Please amend the subtitle beginning on page 3, line 20 as follows:

~~Best Mode for Carrying Out~~ Detailed Description of the Invention

Please amend the paragraph beginning on page 9, line 16 with the following amended paragraph:

Server 1 ~~analyses~~analyzes the request information transmitted from client 11 by a request analyzer 4 (request information analyzing means (including request information receiving means)) and selects the necessary image data from multiple viewpoint image data 2 (multiple viewpoint image supply means) to output it to an image generator 3 (image generating means) where image data for the requested viewpoint (viewpoint information) is generated by interpolation to be output to an image synthesizer 5 (image synthesizing means). In image synthesizer 5, a plurality of supplied images data are synthesized in a form (format based on the display unit information) suitable for encoding to be output to an encoder 6 (coding means). In encoder 6, the supplied image data is encoded at a suitable bit rate to be transmitted to network 7 (transmitting means).

Please amend the paragraph beginning on page 10, line 6 with the following amended paragraph:

Client 11 receives the coded image data (via receiving means~~means~~ 10), and decodes the data through a decoder 12 (decoding means) and outputs the decoded image data to an image processor 13 (image processing means), where the image data is converted into an appropriate form in conformity with a stereoscopic display format so that the image data is displayed on a display unit 14 (display means). Client 11 also includes an input unit 16 (request information input means) for change of the viewpoint, and transmits the request information of viewpoint alternation to network 7 by way of a request output unit 15 (request information transmitting means).

Please amend the paragraph beginning on page 10, line 18 with the following amended paragraph:

Multiple viewpoint image data 2 is formed of a set of images data taken by a plurality of cameras. The plurality of cameras are typically laid out as shown in ~~Fig.2 (a)~~Fig. 2A so that the optical axes of the plurality of cameras intersect at one point. As a special example, the cameras may be arranged on the

circumference of a circle so that the optical axes of the cameras are directed to the center of the circle, as shown in ~~Fig. 2 (b)~~Fig. 2B. In either case, the cameras are not necessarily arranged equi-distantly, but may be laid out in some parts densely and others sparsely. The information as to how the cameras are allocated is also recorded together with the image data. This allocation information is used to determine the image data of which cameras should be used when image generator 3 generates the image data from a designated viewpoint by interpolation based on the request information from client 11.

Please amend the paragraph beginning on page 11, line 9 with the following amended paragraph:

Next, description will be made about the necessary camera image data when the image of the requested viewpoint is generated by interpolation. In the example shown in ~~Fig. 3 (a)~~Fig. 3A, the left eye viewpoint L and right eye viewpoint R are designated as illustrated with respect to cameras C1 to C4. In this case, the images of data from C1 and C2 are used to generate the image data for left eye viewpoint L and the images of data from C2 and C3 are used to generate the image data for right eye viewpoint R. Similarly, in the example shown in ~~Fig. 3 (b)~~Fig. 3B, the layout of cameras C1 to C4 and the requested left eye viewpoint L and right eye viewpoint R are positioned as illustrated. In this case, both the image data for left eye viewpoint L and the image data for right eye viewpoint R are generated based on the images of data from C1 and C2. Though four cameras are used in the examples in ~~Fig. 3~~ Figs. 3A and 3B, the number of cameras is not limited to four.

Please amend the paragraph beginning on page 12, line 9 with the following amended paragraph:

When the necessary viewpoint image data has been generated at image generator 3, image synthesizer 5 implements an extraction process of image data in an amount for the requested resolution. In ~~Fig. 4 (a)~~Fig. 4A, the size of the generated image data 21 is assumed to be equal to the size of the image taken by the camera, and the resolution required for display on client 11 is shown by an area 22. In ~~Fig. 4~~

~~(a)Fig. 4A~~, only part of the generated image data is cut out. Alternatively, as shown in ~~Fig.4 (b)Fig. 4B~~, instead of cutting out an image in the resolution required for display on client 11, the maximum area 24 capable of being cut out while keeping the display aspect may be cut out from the generated image data 23 (which is assumed to be the same size as the image taken by the camera), then may be reduced to the required resolution. Also, image generator 3 may be adapted to generate only the necessary amount of image data for the required resolution.

Please amend the paragraph beginning on page 13, line 7 with the following amended paragraph:

~~Fig.6~~Figs. 6A, 6B, and 6C ~~show~~shows examples of storage states of multiple viewpoint image data 2. Because moving picture data taken by a plurality of cameras need to be temporally synchronized to each other, the plurality of images data C1 to C4 taken by different cameras may be stored in such a manner as to join them abreast and form a piece of image data made of one image, as shown in ~~Fig-6~~(a)Fig. 6A. This format assures that the images data C1 to C4 contained in the image data of the single image were taken at the same time, leading to easy time management. The way of joining is not limited to a line abreast as shown in ~~Fig-6 (a)Fig. 6A~~ but the images may be joined as shown in ~~Fig-6 (b)Fig. 6B~~, for example.

Please amend the paragraph beginning on page 13, line 19 with the following amended paragraph:

In contrast, as shown in ~~Fig-6 (e)Fig. 6C~~, a format of separately storing the images of data C1 to C4 taken by different cameras may be also considered. The advantage of this method is that, when camera images C1 and C2 alone are needed as shown in ~~Fig-3 (b)Fig. 3B~~ in order to generate the image data from the requested viewpoint, these can be easily picked up.

Please amend the paragraph beginning on page 13, line 25 with the following amended paragraph:

Multiple viewpoint image data 2 may be stored either by being compressed or non-compressed. Here, referring to ~~Fig. 8~~Figs. 8A and 8B, description will be made about a case where individual pieces of camera image data which are given separately in the manner as shown in ~~Fig. 6 (e)~~Fig. 6C are stored after being data compressed. In this case, images of data C1 to C4 taken by the individual cameras are input to an encoder 31 as shown in ~~Fig. 8 (a)~~Fig. 8A. Encoder 31 encodes each image data and outputs the necessary information (frame type, generation bit count, etc.) for generating management information, to a management information generator 32. In a recorder 33, the coded image data and the management information are recorded as the storage data (management information adding means). The detail of the management information and storage format will be described later.

Please amend the paragraph beginning on page 14, line 15 with the following amended paragraph:

When the storage data which has been recorded in the manner shown in ~~Fig. 8 (a)~~Fig. 8A is used as the multiple viewpoint image data, the data needs to be decoded so as to be handled by image generator 3. In this case, as shown in ~~Fig. 8 (b)~~Fig. 8B, a selector 34, in accordance with the request from the client, selects only the necessary image data from the stored data and outputs to a decoder 35 where the data is decoded, whereby it is possible to obtain the necessary and sufficient original image data (the image data to be used by image generator 3). In this process, in order to extract the necessary part quickly, the management information recorded together with image data is utilized.

Please amend the paragraph beginning on page 15, line 2 with the following amended paragraph:

~~Fig. 7~~Figs. 7A, 7B, and 7C ~~show~~shows examples of image data states generated by image synthesizer 5. Though the left eye viewpoint image data L and the right eye viewpoint image data R can be encoded as separate pieces of image data as shown in ~~Fig. 7 (b)~~Fig. 7B, it is preferred that both images of data are joined side by side (or up and down) into a piece of synthesized image data formed of a single

image then the joined image data is encoded. Formation of the image data composed of a single image as shown in ~~Fig. 7 (a)~~Fig. 7A is able to assure the synchronism between the left eye viewpoint image data L and the right eye viewpoint image data R, leading to easy time management. Particularly, when a coding scheme entailing frame skipping such as MPEG-4 is used, it is possible to prevent occurrence of such a situation where, at a certain moment, the frame of the left eye viewpoint image data L is present whereas the frame of the right eye viewpoint image data R is not present. Further, since the image data is rate-controlled as a single image, both the left eye viewpoint image L and the right viewpoint image R can be kept substantially equal in image quality. When the images are encoded as separate pieces of image data, depending on the rate control result there occurs a case where on extraction of a frame at a certain moment the image quality of the left eye viewpoint image L is good while the image quality of the right eye viewpoint image R is bad. In such a case, the resultant stereoscopic display presents poor quality. Thus, it is possible to improve the quality of stereoscopic display when the data is adapted to take the form as shown in ~~Fig. 7 (a)~~Fig. 7A.

Please amend the paragraph beginning on page 16, line 5 with the following amended paragraph:

Depending on the stereoscopic display format of a client 11, there are some cases where the image data to be finally displayed on the display unit 14 takes a form of strips of left eye viewpoint image data L and strips of right eye viewpoint image data R being alternated with each other every line as shown in ~~Fig. 7 (e)~~Fig. 7C (for lenticular mode, parallax barrier mode and the like). In such a case, however, it is preferred that the image data to be coded takes the form as shown in ~~Fig. 7 (a)~~Fig. 7A. The reason is that if coding is performed on a block basis as in DCT, the image data having the form as shown in ~~Fig. 7 (e)~~Fig. 7C will present weak correlation between adjacent pixels, hence high spatial frequencies, producing poor compression efficiency. The same method can be applied for the cases where the number of viewpoints is greater than 2.

Please amend the paragraph beginning on page 16, line 19 with the following amended paragraph:

When images of data for a plurality of viewpoints are joined to form a piece of image data made of a single image as shown in ~~Fig. 7 (a)~~Fig. 7A, it is impossible to know the difference between normal two-dimensional image data and stereoscopic image data from a format point of view. No problem will occur when real-time streaming is handled using the system shown in Fig. 1 because image data is transmitted in real time in accordance with the request from the client side. However, in a case where the thus transmitted image data has been once recorded locally on the client 11 side and is played afterward, it is impossible to distinguish whether it is of two-dimensional image data or stereoscopic image data. In order to prevent this, when any piece of image data having the form as shown in ~~Fig. 7 (a)~~Fig. 7A is recorded, a flag for identification indicating whether the image data is of two-dimensional image data or stereoscopic image data may and should be added (identification information adding means). This addition of the identification flag may be done either on server 1 or client 11. Further, client 11 has a ~~judgement~~judgment means for distinction between two-dimensional image data or stereoscopic image data.

Please amend the paragraph beginning on page 18, line 5 with the following amended paragraph:

Fig. 10 is a flowchart showing the sequence of processing on client 11. First, initialization is done so that the viewpoint position (viewpoint information) at the initial state and information not dependent on the viewpoint (stereoscopic display format, resolution etc. (display unit information)) are set up (Step S11). Next, these pieces of information are transmitted as a request to server 1 (Step S12). Then, the bit stream (image data) meeting the request is transmitted from the server by way of the network (Step S13). Subsequently, the bit stream is decoded (Step S14). Since, as shown in ~~Fig. 7 (a)~~Fig. 7A the decoded image data is not in the form that is directly and stereoscopically displayable, the data is rearranged so as



to conform with the stereoscopic display format, as shown in ~~Fig. 7 (e)~~Fig. 7C (Step S15). Then the data is displayed on display unit 14 (Step S16). Next, it is judged whether there is a next display (Step S17). If display is continued, it is judged whether there is a request for change of the viewpoint (Step S18). Subsequently, if the viewpoint should be changed, the request is sent again to server 1 and the operation returns to Step S12. When no more display is needed at Step S18, the operation goes to Step S13.

Please amend the paragraph beginning on page 20, line 17 with the following amended paragraph:

~~Fig. 13 is a chart~~Figs. 13A and 13B are charts showing one example of management information added at management information generator 32 shown in ~~Fig. 8 (a)~~Fig. 8A. The coded data of individual camera images can be joined to management information as shown in Fig. 12 and stored. This management information is the information that allows for access to each camera image data. In the case of a multiple viewpoint moving picture, the management information contains not only the information which allows for access to each camera image but also the information which allows random access to the coded data at a designated time within each camera image data.

Please amend the paragraph beginning on page 21, line 3 with the following amended paragraph:

~~Fig. 13 (a)~~Fig. 13A shows one example of management information for access to the coded data of individual camera images. For example, it is indicated that the coded data of camera image C2 exists at the B2-th byte from the front of the data in Fig. 12. ~~Fig. 13 (a)~~Fig. 13A further includes the pointers to the information for making access to the coded data at designated times within the camera image data. For the coded data of C2, it is indicated that the access table for making access to the coded data at designated times is located at address P2 within the management information.

Please amend the paragraph beginning on page 21, line 13 with the following amended paragraph:

~~Fig. 13 (b)~~Fig. 13B shows one example of an access table to coded data at designated times. Times  $t_1$ ,  $t_2$ ,  $t_3$ , ... may be set up at regular intervals or may be set up arbitrary intervals apart. For example, it is indicated that the coded data corresponding to time  $t_3$  exists at the  $Bt_3$ -th byte from the front of the coded data of the camera image while the coded data of I-frame is located at a position upstream by  $It_3$  bytes from the aforementioned position. If the decoder needs to start display from time  $t_3$ , the coded data of the I-frame located at the  $(Bt_3 - It_3)$ -th byte from the front is decoded first. Then, P-frames and B-frames are successively decoded while counting the number of bytes of the decoded data until the count reaches  $It_3$  bytes. When the display is started at this point of time, the image data from the designated time  $t_3$  is displayed.

Please amend the paragraph beginning on page 21, line 13 with the following amended paragraph:

Please amend the paragraph beginning on page 22, line 3 with the following amended paragraph:

Next, other accessing methods will be described.

(A) Coded data is packetized and the header information of each packet has information that indicates whether the packet contains the front of an I-frame. In ~~Fig. 13 (b)~~Fig. 13B, the designated time and the number of bytes to the packet corresponding to the designated time are written. When the decoder had made access to the packet corresponding to the designated time  $t_3$ , it is checked as to whether the packet contains the front of an I-frame. The decoder starts decoding and display from a packet that contains an I-frame (all the packets before that are discarded as unnecessary).

(B) In (A), instead of indicating the number of bytes to the packet, only the packet number is written in ~~Fig. 13 (b)~~Fig. 13B. The length of the packets in a piece of coded data is assumed to be fixed and the number of bytes of one packet is written in the header information of the coded data. The decoder calculates the number of bytes to the packet corresponding to the designated time based on the packet

number and the number of bytes of one packet (thereafter the steps are the same as (A)).

Please amend the paragraph beginning on page 22, line 23 with the following amended paragraph:

Next, other storage forms will be described.

(C) In Fig.12, the management information and the coded information are joined and stored.

However, the management information may be separated and stored as a different file.

(D) Of the management information, the information for access to designated times may be included in the header information of the coded data of each camera image, instead of being included in the management information. In this case, the third column in ~~Fig.13 (a)~~Fig. 13A (the pointer to the information for access to designated times within each camera image) is not necessary.

(E) The management information, the coded data of individual camera images may be all separated into different files. In this case, the number of bytes from the front in the second column in ~~Fig.13 (a)~~Fig. 13A is replaced by the filename of the coded data of each camera image, for example. Access to each camera image is made based on the filename.